Technische Universität Ilmenau
Institut für Mathematik

Preprint No. M 98/28

Zur Konstruktion von Differenzenverfahren 2. Ordnung für quasilineare hyperbolische Systeme auf dem Torus

Vogt, Werner

Dezember 1998

Impressum:
Hrsg.: Leiter des Instituts für Mathematik
Weimarer Straße 25
98693 Ilmenau
Tel.: +49 3677 69 3621
Fax: +49 3677 69 3270
http://www.tu-ilmenau.de/ifm/

ISSN xxxx-xxxx

ilmedia

Zur Konstruktion von Differenzenverfahren 2. Ordnung für quasilineare hyperbolische Systeme auf dem Torus

Werner Vogt
Technische Universität Ilmenau
Institut für Mathematik
Postfach 100565
98684 Ilmenau

17. Dezember 1998

Zusammenfassung

Für quasilineare hyperbolische Differentialgleichungssysteme auf dem Standardtorus werden stabile Differenzenverfahren 2. Ordnung mittels globaler Extrapolation und alternativ dazu mittels des Defektkorrektur-Prinzips begründet. Beide Zugänge sind sowohl auf explizite als auch auf linear-implizite Verfahren vom Upwind-Typ anwendbar. Insbesondere ist mittels eines einzigen Nachiterationsschrittes auf der Basislösung 1. Ordnung eine Approximation der Konvergenzordnung 2 sowie eine asymptotische Schätzung des globalen Diskretisierungsfehlers ohne Gitterverfeinerung erreichbar.

1 Einführung

Wir betrachten als Grundgebiet der Differentialgleichungen den p-dimensionalen Standardtorus

$$\mathbb{T}^p = \{\theta \mid \theta = (\theta_1, \cdots, \theta_p)\ ,\ \theta_i = \mathbb{R}\ mod\ 2\pi\ ,\ i = 1(1)p\} \tag{1}$$

und eine darauf definierte Funktion $u : \mathbb{T}^p \to \mathbb{R}^q$. Die Ermittlung eines durch $u(\theta) = (u_1(\theta), \cdots, u_p(\theta))$ parametrisierten invarianten Torus führt auf die Torusgleichung

$$\sum_{j=1}^{p} \omega_j(\theta, u) \frac{\partial u}{\partial \theta_j} \ = \ f(\theta, u)\ , \quad \theta \in \mathbb{T}^p \tag{2}$$

mit den Torus-Bedingungen

$$u(\theta_1, ..., \theta_{k-1}, \theta_k, \theta_{k+1}, ..., \theta_p) = u(\theta_1, ..., \theta_{k-1}, \theta_k + 2\pi, \theta_{k+1}, ..., \theta_p),\ k = 1(1)p. \tag{3}$$

Dieses quasilineare System mit gleichem Hauptteil wird für den speziellen Fall des 2-Torus mit $\omega_1(\theta, u) \neq 0$ auf $\mathbb{T}^2 \times \mathbb{R}^q$ in [1] mit der Darstellung als zeitabhängiges Problem in $(t, \theta) \in \mathbb{T}^2$ behandelt und mit stabilen Differenzenverfahren 1. Ordnung gelöst. Um hinreichende Genauigkeit zu erzielen, ist allerdings auf einem sehr feinen Gitter

$$\mathbb{T}_h^2 = \{(\theta_j, t_h) \mid t_h = n \cdot \tau,\ \theta_j = j \cdot h,\ n = 0(1)N,\ j = 0(1)J\} \tag{4}$$

mit $\tau = 2\pi/N$ und $h = 2\pi/J$ zu approximieren, was zu beträchtlichen Rechenzeiten führen kann. Zugleich liefern die Verfahren weder eine Schätzung des lokalen noch des globalen Diskretisierungsfehlers der Lösung der Randwertaufgabe (2), (3).

Um Approximationen höherer Ordnung und gleichzeitig asymptotische Fehlerschätzungen zu gewinnen, bieten sich insbesondere Extrapolationsverfahren und Defektkorrektur-Verfahren (deferred correction methods) an. Während die Richardson-Extrapolation gegenwärtig auch für partielle DGL-Systeme (vgl. [2]) häufiger genutzt wird, sind Defektkorrektur-Verfahren (vgl. [3]) weniger bekannt - vermutlich wegen ihres komplizierten theoretischen Hintergrunds.

Nachfolgend wird ein einheitlicher Zugang für alle betrachteten - expliziten und impliziten - Basis-Verfahren 1. Ordnung angegeben, der konvergente Lösungen 2. Ordnung und asymptotische Schätzungen des globalen Fehlers der Basislösung liefert. Im Unterschied zur Extrapolation ist bei der Defektkorrektur keine Gitterverfeinerung nötig. Werden die nichtlinearen Gleichungssysteme zudem mit einem Newton-ähnlichen Verfahren gelöst, so ist lediglich *ein zusätzlicher Newtonschritt* mit modifizierter rechter Seite erforderlich.
Im folgenden Abschnitt 2 werden summarisch die Basisverfahren 1. Ordnung verifiziert, auf denen die Approximationen 2. Ordnung aufbauen. Während die Richardson-

Extrapolation in Abschnitt 3 exemplarisch vorgestellt wird, wird die Defektkorrektur nach einer theoretischen Begründung (Abschnitt 4) ausführlicher in Abschnitt 5 behandelt. Ein Anwendungsbeispiel demonstriert schließlich den Wert der Verfahren 2. Ordnung.

2 Basisdiskretisierung auf dem Standardtorus

Sei B^0 der Banachraum $C^0(\mathbb{T}^2, \mathbb{R}^q)$ der auf $\mathbb{T}^2$ stetigen Funktionen v mit der Norm

$$||v||_0 = \max_{\mathbb{T}^2} ||v(t,\theta)||_\infty \, .$$

Wir betrachten nachfolgend die auf $\mathbb{T}^2$ stetig differenzierbaren Funktionen (die damit bezüglich θ und t 2π-periodisch sind), und definieren mit der Norm

$$\|u\|_1 = max\{\|u\|_0, \|u_t\|_0, \|u_\theta\|_0\}$$

den Banachraum

$$B = \{u | u \in C^1(\mathbb{T}^2, \mathbb{R}^q),\ u(\theta,t) = u(\theta + 2\pi, t) = u(\theta, t + 2\pi),\ (\theta,t) \in \mathbb{T}^2\} \, . \tag{5}$$

Das Torusproblem (2), (3) für den 2-dimensionalen Fall lautet mit $u \in B$ nunmehr

$$u_t + \omega(\theta,t,u)u_\theta = f(\theta,t,u) \, , \tag{6}$$

wobei die Funktionen $\omega : \mathbb{T}^2 \times \mathbb{R}^q \to \mathbb{R}$ und $f : \mathbb{T}^2 \times \mathbb{R}^q \to \mathbb{R}^q$ als hinreichend glatt vorausgesetzt werden. Definiert man den Operator $F : B \to B^0$ durch

$$(Fu)(\theta,t) \ \equiv \ u_t(\theta,t) + \omega(\theta,t,u(\theta,t))u_\theta(\theta,t) - f(\theta,t,u(\theta,t)), \tag{7}$$

so lautet das Problem in Operatorschreibweise

$$Fu = 0, \quad u \in B \, . \tag{8}$$

Um die betrachteten Differenzenverfahren in eine analoge Operatorform zu überführen, betrachten wir auf dem diskretisierten Torus $\mathbb{T}_h^2$ gemäß (4) mit der Schrittweitenbedingung

$$\lambda := \tau/h = const \tag{9}$$

entsprechende Banach-Räume B_h und B_h^0 von Gitterfunktionen $u_h = \{u_j^u\}$
mit $j = 0(1)J-1,\ n = 0(1)N-1$. Dabei approximiere $u_j^n \sim u(\theta_j, t_n)$ auf $\mathbb{T}_h^2$.

Sei $\|u_j^u\|_\infty$ die Maximumnorm des $\mathbb{R}^q$, so ist

$$\|u_h\|_0 = \max_{\mathbb{T}_h^2} \ \|u_j^n\|_\infty$$

die diskrete C-Norm und $B_h^0 = C_h^0(\mathbb{T}_h^2, \mathbb{R}^q)$ der entsprechende Banach-Raum. Bezeichnet man mit $\partial_t u$ und $\partial_\theta u$ die Differenzenquotienten

$$\begin{aligned} \{\partial_t u_h\}_j^n &= \frac{1}{\tau}(u_j^n - u_j^{n-1}), \\ \{\partial_\theta u_h\}_j^u &= \frac{1}{h}(u_j^n - u_{j-1}^n), \end{aligned} \tag{10}$$

so erhält man die diskrete C^1-Norm

$$\|u_h\|_1 = max\{\|u_h\|_0, \|\partial_t u_k\|_0, \|\partial_\theta u_h\|_0\}$$

und damit den Banach-Raum

$$B_h = \{u_h | u_h \in C_h^1(T_h^2, \mathbb{R}^q),\ u_j^n = u_{j\ mod\ J}^n = u_j^{n\ mod\ N},\ u_h \in \mathbb{T}_h^2\}\ . \tag{11}$$

Der Operator $F_h : B_h \to B_h^0$ ist allgemein in der Form des 6-Punkt- Schemas

$$\{F_h u_h\}_j^n \equiv \frac{1}{\tau}\left\{\sum_{\mu=-1}^{1} S_\mu^*(\theta_j, t_h, u_j^n) u_{j+\mu}^{n+1} - \sum_{\mu=-1}^{1} S_\mu(\theta_j, t_h, u_j^n) u_{j+\mu}^n\right\} - f(\theta_j, t_h, u_j^n) \tag{12}$$

mit $j = 0(1)J-1$, $n = 0(1)N-1$ definiert. $S_\mu^*(\theta_j, t_h, u_j^n)$ und $S_\mu(\theta_j, t_h, u_j^n)$ sind q-reihige Diagonalmatrizen für $\mu = -1, 0, 1$. Mit F_h lauten die betrachteten Verfahren in Operatorschreibweise

$$F_h u_h = 0, \quad u_h \in B_h\ . \tag{13}$$

In der allgemeinen Operatordarstellung (12) sind alle nachfolgend behandelten expliziten und linear-impliziten 6-Punkt-Schemata enthalten.

1. Explizite Verfahren vom Upwind–Typ

Bei diesen Verfahren wird die Matrix $C(\theta_j, t_n, u_j^n)$ in die Summe einer Matrix $C^+(\theta_j, t_n, u_j^n)$ mit nur positiven Elementen und einer Matrix $C^-(\theta_j, t_n, u_j^n)$ mit nur negativen Elementen zerlegt :

$$C(\theta_j, t_n, u_j^n) = C^+(\theta_j, t_n, u_j^n) + C^-(\theta_j, t_n, u_j^n) \quad . \tag{14}$$

(a) Das *explizite Courant-Isaacson-Rees Verfahren* (CIR) ergibt sich mit

$$\begin{aligned} C^+ &= diag\left(\frac{1}{2}(c_i + \mid c_i \mid)\right) \geq 0 \\ C^- &= diag\left(\frac{1}{2}(c_i - \mid c_i \mid)\right) \leq 0 \quad , \end{aligned} \tag{15}$$

(b) das *explizite glatte Upwind-Verfahren* mit der Wahl

$$\begin{aligned} C^+ &= diag\left(\frac{1}{2}(c_i + \Phi(c_i))\right) \geq 0 \\ C^- &= diag\left(\frac{1}{2}(c_i - \Phi(c_i))\right) \leq 0 \quad , \\ \text{mit} \quad \Phi(d) &:= \sqrt{\delta^2 + d^2} \quad , \quad \delta \neq 0 \ , \ constant \ . \end{aligned} \tag{16}$$

Die Diagonalmatrizen S_μ and S_μ^* , $\mu = -1, 0, 1$, besitzen für alle expliziten Verfahren vom Upwind-Typ die Form

$$\begin{aligned} S_{-1} &= \lambda C^+ \quad , \quad S_{-1}^* = 0 \\ S_0 &= I - \lambda C^+ + \lambda C^- \quad , \quad S_0^* = I \\ S_1 &= -\lambda C^- \quad , \quad S_1^* = 0 \ . \end{aligned} \tag{17}$$

In der üblichen Notation läßt sich das explizite Upwind-Verfahren als

$$\begin{aligned} \{F_h u_h\}_j^n \equiv & \frac{1}{\tau}(u_j^{n+1} - u_j^n) + \omega(\theta_j, t_n u_j^n)\frac{1}{2h}(u_{j+1}^n - u_{j-1}^n) - \\ & -\Phi(\omega(\theta_j, t_n, u_j^n))\frac{1}{2h}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) - \\ & -r(\theta_j, t_n, u_j^n) = 0 \quad \text{mit} \quad (\theta_j, t_n) \in \mathbb{T}_h^2 \\ & \text{und} \quad \Phi(d) = \sqrt{\delta^2 + d^2}, \delta \neq 0, const. \end{aligned} \tag{18}$$

notieren.

2. Explizites Friedrichs–Verfahren

In diesem Falle wird die Matrix $C(\theta_j, t_n, u_j^n)$ nicht zerlegt, und die Diagonalmatrizen S_μ and S_μ^* , $\mu = -1, 0, 1$ besitzen nun die Form

$$\begin{aligned} S_{-1} &= \frac{1}{2}(I + \lambda C) \quad , \quad S_{-1}^* = 0 \\ S_0 &= 0 \quad , \quad S_0^* = I \\ S_1 &= \frac{1}{2}(I - \lambda C) \quad , \quad S_1^* = 0 \quad . \end{aligned} \tag{19}$$

3. Linear implizite Verfahren

Analog zu den expliziten Verfahren vom Upwind-Typ wird die Matrix $C(\theta_j, t_n, u_j^n)$ in die beiden Matrizen $C^+(\theta_j, t_n, u_j^n)$ und $C^-(\theta_j, t_n, u_j^n)$ zerlegt, wobei im Falle des impliziten CIR-Verfahrens

$$\begin{aligned} C^+ &= diag\left(\frac{1}{2}(c_i + \mid c_i \mid)\right) \geq 0 \\ C^- &= diag\left(\frac{1}{2}(c_i - \mid c_i \mid)\right) \leq 0 \quad , \end{aligned} \tag{20}$$

und im Falle des impliziten glatten Upwind-Verfahrens

$$\begin{aligned} C^+ &= diag\left(\frac{1}{2}(c_i + \Phi(c_i))\right) \geq 0 \\ C^- &= diag\left(\frac{1}{2}(c_i - \Phi(c_i))\right) \leq 0 \quad , \\ \text{with} \quad \Phi(d) &:= \sqrt{\delta^2 + d^2} \quad , \quad \delta \neq 0 \, , \, constant \end{aligned} \tag{21}$$

gilt. Für alle derartigen linear impliziten Verfahren sind die Diagonalmatrizen S_μ und S_μ^* , $\mu = -1, 0, 1$ durch

$$\begin{aligned} S_1 = 0 \quad &, \quad S_{-1}^* = -\lambda C^+ \\ S_0 = I \quad &, \quad S_0^* = I - \lambda C^- + \lambda C^+ \\ S_1 = 0 \quad &, \quad S_1^* = \lambda C^- \quad . \end{aligned} \tag{22}$$

gegeben.

In der Arbeit [6] wird die diskrete Konvergenz der Verfahren (13) mit Operatoren (12) nachgewiesen. Da nachfolgend auf die dortigen Voraussetzungen Bezug genommen wird, sollen die entsprechenden 2 Sätze zitiert werden.

Satz 1 *Der Differenzenoperator (12) für (6) genüge folgenden Voraussetzungen:*

(i) $\lambda := \dfrac{\tau}{h} = const$

(ii) $\omega(\theta, t, u) \in \mathcal{C}^2(\mathbb{T}^2 \times \mathbb{R}^q, \mathbb{R}), f(\theta, t, u) \in \mathcal{C}^2(\mathbb{T}^2 \times \mathbb{R}^q, \mathbb{R}^q)$

(iii) *(8) besitzt eine lokal eindeutige Lösung* $u \in \mathcal{C}^2(\mathbb{T}^2)$

(iv) $\sum\limits_{\mu=-1}^{1} S_\mu(\theta, t, u) \equiv I \quad , \quad \sum\limits_{\mu=-1}^{1} S_\mu^*(\theta, t, u) \equiv I$

(v) $\sum\limits_{\mu=-1}^{1} \mu(S_\mu^*(\theta, t, u) - S_\mu(\theta, t, u)) \equiv \lambda C \;\; where \;\; C = (c_{ij})$

$$c_{ij} = \begin{cases} \psi(\theta, t, u) & f\ddot{u}r \; i = j \\ 0 & sonst \; . \end{cases}$$

Dann ist jedes Verfahren (13) konsistent in h *und* τ *mit Ordnung 1. Gilt desweiteren*

(vi) *(13) ist von positivem Typ, die Diagonalmatrizen* S_μ^* , $\mu = -1, 0, 1$ *sind Lipschitz-stetig in* u *und zwei der drei Matrizen* S_μ^* *besitzen Elemente* $(s_\mu^*)_{kl} \leq 0$, $k, l = 1(1)q$

(vii) *Für die Anfangswerte sei* $\| u_j^0 - g(\theta_j) \| \leq K_0 h \;\; \forall j$, $h \leq h_0$, $K_0 > 0$,

so konvergiert jedes Verfahren (13) mit Ordnung 1 auf $\mathbb{T}_h^2$, *d.h.*

$$\| e_j^n \| = \| u_j^n - u(\theta_j, t_n) \|_\infty \leq Kh \; , \; K > 0 \quad \forall j, n \; .$$

Beweis: Vgl. [6] , S. 7 - 12.

Ein Differenzenoperator (12) ist dabei von *positivem Typ*, falls die 3 Matrizen $S_\mu(\theta,t,u)$, $\mu = -1,0,1$, für alle $(\theta,t,u) \in \mathbb{T}^2 \times \mathbb{R}^q$, $\| u \| \leq M$, M const, nur nichtnegative Elemente besitzen.

Im Falle der drei *expliziten Verfahren* sind die Voraussetzungen (iv) und (v) erfüllt. Der folgende Satz liefert eine hinreichende Bedingung für die Positivität (vi) der Verfahren.

Satz 2 *Sei auf der Menge* $G = \{(\theta,t,u) \mid (\theta,t) \in \mathbb{T}^2, \| u \| \leq M, M \in \mathbb{R}, const\}$ *die Schrittweitenbedingung*

$$\lambda \leq \frac{1}{D} \quad mit \quad D := \max_G \max_{i=1(1)q} \mid c_i(\theta,t,u) \mid \quad . \tag{23}$$

gegeben. Dann sind das explizite CIR-Verfahren und das explizite Friedrichs-Verfahren von positivem Typ. Ist jedoch

$$\lambda \leq \frac{1}{\sqrt{\delta^2 + D^2}} , \tag{24}$$

so ist auch das explizite glatte Upwind-Verfahren von positivem Typ.

Beweis: Vgl. [6] , S. 13.

Die beiden *linear impliziten Verfahren* sind von positivem Typ für alle $\lambda \in \mathbb{R}^+$, und die Diagonalmatrizen S^*_μ , $\mu = -1,0,1$, sind Lipschitz-stetig in u. Desweiteren ergibt sich $S^*_{-1} \leq 0$ and $S^*_1 \leq 0$ aus (20), (21) und (22). Damit sind beide Verfahren konsistent unter den Voraussetzungen (i)- (iv) und konvergent bei geeigneten Anfangswerten.

3 Globale Extrapolation und Fehlerschätzung

Für die Anwendung von Verfahren höherer Konvergenzordnung besitzt die Aufgabe (2), (3) gute Voraussetzungen, da die Koeffizientenfunktionen ω und r sowie die Lösung u als hinreichend glatt vorausgesetzt werden können und Randapproximationen des Integrationsgebietes nicht erforderlich sind.
Um zu einem Verfahren der Ordnung 2 zu gelangen, bietet sich insbesondere die Richardson-Extrapolation an, die zugleich eine asymptotische Schätzung des globalen Diskretisierungsfehlers des Basisverfahrens liefert. Dieses Verfahren bezieht sich auf die zu lösende Randwertaufgabe (2), (3) und ist wesentlich zu unterscheiden von der Lösung der dabei auftretenden Anfangswertaufgaben, wenn z.B. ein Schießverfahren benutzt wird. Auch hierfür sind lokale bzw. globale Extrapolation anwendbar; allerdings ist zu beachten, daß damit noch keine Fehlerschätzung für die zu lösende

Randwertaufgabe (2), (3) gewonnen wird!

Grundlage für die globale Extrapolation bildet der Nachweis der Existenz einer asymptotischen Entwicklung des globalen Diskretisierungsfehlers für das Basisverfahren (12)

$$e_j^n := u_j^n - u(\theta_j, t_n) = e(\theta_j, t_n)h + O(h^2) \tag{25}$$

mit einer hinreichend glatten Funktion $e : \mathbb{T}^2 \to \mathbb{R}^q$, wobei weiterhin $\lambda := \tau/h = const$ vorausgesetzt wird. Dieser Nachweis basiert auf dem bekannten Satz von STETTER (vgl. [3], Theorem 1.3.1.), dessen Anwendung exemplarisch für das explizite glatte Upwind-Verfahren skizziert werden soll.

Sei $B^r = C^r(\mathbb{T}^2, \mathbb{R}^q)$, $r = 1, 2, 3$, der Banachraum aller r-mal stetig differenzierbaren Torusfunktionen mit entsprechender C^r-Norm. Mit dem Operator $F : B^r \to B^{r-1}$

$$(Fu)(\theta, t) \equiv \frac{\partial u}{\partial t} + \omega(\theta, t, u)\frac{\partial u}{\partial \theta} - r(\theta, t, u) \tag{26}$$

läßt sich Aufgabe (2), (3) abkürzend als $Fu = 0$ in B^r notieren. Für das *explizite Upwind-Verfahren* führen wir den Differenzenoperator F_h auf $\mathbb{T}_h^2$ durch

$$\begin{aligned} \{F_h u_h\}_j^n \quad \equiv \quad & \frac{1}{\tau}(u_j^{n+1} - u_j^n) + \omega(\theta_j, t_n u_j^n)\frac{1}{2h}(u_{j+1}^n - u_{j-1}^n) - \\ & -\Phi(\omega(\theta_j, t_n, u_j^n))\frac{1}{2h}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) - \\ & -r(\theta_j, t_n, u_j^n) \;=\; 0 \quad \text{mit} \quad (\theta_j, t_n) \in \mathbb{T}_h^2 \\ & \text{und} \quad \Phi(d) = \sqrt{\delta^2 + d^2}, \delta \neq 0, const, \end{aligned} \tag{27}$$

ein, mit dem das Verfahren $F_h u_h = 0$ in endlichdimensionalen Banachräumen B_h^r, $r = 1, 2, 3$, lautet. Die Existenz der asymptotischen Entwicklung (25) wird garantiert durch

Satz 3 *Das explizite Upwind-Verfahren (27) erfülle die Voraussetzungen der Sätze 1 und 2. Weiterhin gelten folgende Voraussetzungen:*

(i) Die Lösung $u^* \in C^3(\mathbb{T}^2, \mathbb{R}^q)$ sei regulär (d.h. der Operator $F'(u^*)$ besitzt einen beschränkten inversen Operator)

(ii) $\Phi \in C^2(\mathbb{R})$.

Dann existiert eine Torusfunktion $e \in C^2(\mathbb{T}^2, \mathbb{R}^q)$, mit der die asymptotische Entwicklung (25) des globalen Fehlers gilt.

Beweis: Nach Satz (1) sind Konsistenz, Stabilität und damit Konvergenz des Verfahrens gesichert. Nach dem Theorem von STETTER ist außerdem die Abbildung des lokalen Fehlers $G : B^r \to B_h^{r-1}$

$$Gu(\theta_j, t_n) \equiv F_h u(\theta_j, t_n) - (Fu)(\theta_j, t_n) \tag{28}$$

zu untersuchen. Sei $u \in B^3 = C^3(\mathbb{T}^2, \mathbb{R}^q)$. Setzt man (26) und (27) ein und entwickelt die Funktion $u(\theta, t)$ an der Stelle (θ_j, t_n), so erhält man unmittelbar

$$Gu(\theta_j, t_n) = (Lu)(\theta_j, t_n) \cdot h + O(h^2) \tag{29}$$

mit dem Operator $L : B^2 \to B^1$

$$(Lu)(\theta, t) \equiv \frac{1}{2}\left[\lambda \frac{\partial^2 u}{\partial t^2}(\theta, t) - \Phi(\omega(\theta, t, u(\theta, t)))\frac{\partial^2 u}{\partial \theta^2}\right]. \tag{30}$$

Speziell für die Lösung u^* ergibt (29) die asymptotische Entwicklung des lokalen Diskretisierungsfehlers

$$\tau_j^n = (Lu^*)(\theta_j, t_n)h + O(h^2) \tag{31}$$

auf $\mathbb{T}_h^2$. Unter den Voraussetzungen des Satzes existieren die Fréchet-Ableitungen $F_h'(u)$, $F'(u)$ und $L'(u)$ und gestatten die Entwicklung

$$\begin{aligned} G'(u)\varepsilon(\theta_j, t_n) &= F_h'(u)\varepsilon(\theta_j, t_n) - (F'(u)\varepsilon)(\theta_j, t_n) \\ &= (L'(u)\varepsilon)(\theta_j, t_n) \cdot h + O(h^2) \end{aligned}$$

mit L gemäß 30 für $\varepsilon \in B^2$. Damit sind auch die 1. Ableitungen F_h' und F' asymptotisch vertauschbar, womit sämtliche Voraussetzungen des Theorems von STETTER erfüllt sind. Dieses liefert unmittelbar (25). □

Bemerkungen:

1. Für das explizite CIR-Verfahren ist wegen $\Phi(d) = |d|$ die Voraussetzung (ii) des Satzes 3 nicht erfüllt. In vielen Anwendungen ist jedoch $\omega(\theta, t, u^*(\theta, t)) \neq 0$ und damit die Voraussetzung erfüllbar.

2. Das *linear implizite Upwind-Verfahren* erfüllt unter den dortigen Voraussetzungen ebenfalls den Satz 3, allerdings hat der Operator (30) nun die kompliziertere Form

$$(Lu)(\theta, t) \equiv \frac{1}{2}\left[\lambda \frac{\partial^2 u}{\partial t^2}(\theta, t) + 2\lambda\Omega \frac{\partial^2 u}{\partial \theta \partial t}(\theta, t) - \Phi(\Omega)\frac{\partial^2 u}{\partial \theta^2}(\theta, t)\right] \tag{32}$$

mit $\Omega := \omega(\theta, t, u(\theta, t))$.

Auf der Grundlage der asymptotischen Entwicklung (25) läßt sich nunmehr die Richardson-Extrapolation anwenden. Seien zwei Gitter $\mathbb{T}^2_{h_i},\ i = 1,2$ mit Schrittweiten (h_i, τ_i) gegeben, die der Relation $\tau_i = \lambda \cdot h_i,\ i = 1,2$, mit der Konstanten $\lambda > 0$ genügen. Zudem kann $h_1 > h_2 > 0$ vorausgesetzt werden. Die Lösungen auf den Gittern $\mathbb{T}^2_{h_i}$ seien $u_{h_i},\ i = 1,2$. Auf dem Gitterdurchschnitt $\mathbb{T}^2_h = \mathbb{T}^2_{h_1} \cap \mathbb{T}^2_{h_2}$ erhält man die extrapolierte Lösung v_h mittels

$$v_h = u_{h_2} - \beta(u_{h_1} - u_{h_2}) \tag{33}$$

mit $\beta = 1/(\frac{h_1}{h_2} - 1)$. Durch Einsetzen der Entwicklung (25) zeigt man unmittelbar

Satz 4 *Unter den Voraussetzungen des Satzes 3 gilt*

(i) $\{v_h\}^n_j = u(\theta_j, t_n) + O(h^2)$ *auf* $\mathbb{T}^2_h$

(ii) $e_{h_1} := \beta(u_{h_1} - u_{h_2})$ *ist eine asymptotische Schätzung des globalen Fehlers der Lösung* u_{h_1}.

Zur algorithmischen Realisierung von (33) wird $h_2 := h$ und $h_1 := 2h$ gesetzt, womit $\beta = 1$ und $e_{2h} := u_{2h} - u_h$ ist.

Algorithmus 1 (Globale Extrapolation)

S_1: Lösung der Randwertaufgabe $F_{2h}u_{2h} = 0$ auf dem Grobgitter $\mathbb{T}^2_{2h}$ liefert u_{2h}.

S_2: Bereitstellung der Startlösung $\{u_h\}^0_j$ auf dem Feingitter $\mathbb{T}^2_h$ durch Interpolation von u_{2h}.

S_3: Lösung der Aufgabe $F_h u_h = 0$ auf dem Feingitter liefert u_h.

S_4: Extrapolation $v_{2h} := u_h - (u_{2h} - u_h)$ auf dem Grobgitter $\mathbb{T}^2_{2h}$ sowie Fehlerschätzung $e_{2h} := u_{2h} - u_h$.

Zur Lösung der Randwertaufgaben in den Schritten S_1 und S_3 bietet sich ein Schießverfahren an: Man gebe eine $2\pi-$periodische Anfangslösung für $t = 0$ vor und löse die zugehörige Anfangswertaufgabe bezüglich der Zeit t so, daß auf jeder Zeitschicht t die Lösung periodisch bleibt.

Um während des Schießverfahrens bereits über eine extrapolierte Lösung und eine globale Fehlerschätzung (für die Anfangswertaufgaben!) zu verfügen, wird man zu einer Startlösung $\{u_h\}^0_j$ durch Restriktion eine Startlösung $\{u_{2h}\}^0_j$ auf dem Grobgitter ermitteln. Die Schritte S_1 und S_3 kann man nun auf jeder Zeitschicht des Grobgitters $\mathbb{T}^2_{2h}$ unabhängig ausführen und die Extrapolation (S_4) schrittweise anwenden. Allerdings liefert $e_{2h} = u_{2h} - u_h$ dann nicht den globalen Fehler der Randwertaufgabe. Dazu sollte man nach Erhalt von u_{2h} mit einigen Zusatziterationen gemäß Schritt S_3 die Lösung u_h und schließlich e_{2h} und v_{2h} bestimmen.

4 Defektkorrektur und Fehlerschätzung

Es wird im folgenden vorausgesetzt, daß das kontinuierliche Problem (8) eine lokal eindeutige Lösung $u^* \in B$ besitzt. Falls das zugehörige diskrete Problem (13) eine Gitterlösung besitzt - dies muß nicht vorausgesetzt werden -, so werde diese "Basislösung" mit $u_h^* \in B_h$ bezeichnet. Die Restriktionsoperatoren p_h mit

$$\{p_h u\}_j^n \equiv u(\theta_j, t_h), \quad j = 0(1)J-1, n = 0(1)N-1 \tag{34}$$

bilden $p_h : B \to B_h$ und $p_h : B^0 \to B_h^0$ ab. Einsetzen der exakten Lösung u^* in das diskrete System liefert den lokalen Diskretisierungsfehler

$$\tau_h \equiv F_h p_h u^* - p_h F u^*. \tag{35}$$

Könnte man τ_h exakt berechnen, so ließe sich damit die gestörte Gleichung

$$F_h v_h = \tau_h, \quad v_h \in B_h \tag{36}$$

aufstellen, deren Lösung nach Definition (35) genau $v_h = p_h u^*$, d. h. das Bild der exakten Lösung wäre. Die Grundidee der Defektkorrektur (DC) besteht nun darin, eine hinreichend genaue Approximation des lokalen Diskretisierungsfehlers τ_h unter Benutzung der Basislösung u_h^* vorzunehmen und mit (36) eine korrigierte Lösung v_h^* zu ermitteln, die die exakte Lösung $p_h u^*$ besser als die Basislösung u_h^* approximiert. Da anstelle der exakten Lösung $u^* \in B$ lediglich die diskrete Lösung $u_h^* \in B_h$ verfügbar ist, muß der letzte Term $p_h F y^*$ des lokalen Fehlers (35) durch einen defektdefinierenden Operator F_h^1 approximiert werden, so daß (35) durch die Näherung

$$l_h = F_h u_h - F_h^1 u_h, \quad u_h \in B_h \tag{37}$$

ersetzt wird. Speziell mit u_h^* ergibt sich hiermit die Darstellung

$$l_h = -F_h^1 u_h^*, \tag{38}$$

die auf folgendes allgemeines DC-Verfahren führt:

Algorithmus 2 (Defektkorrektur)

S_1: Lösung der Randwertaufgabe $F_h u_h = 0$ mit der Basislösung u_h^* .

S_2: Berechnung des Defektes $l_h = -F_h^1 u_h^*$.

S_3: Lösung der korrigierten Gleichung

$$F_h v_h = l_h \tag{39}$$

mit der verbesserten Lösung v_h^* .

Gleichung (39) kann mit demselben stabilen Verfahren gelöst werden, wie die Grundaufgabe (13), wobei eine gute Startnäherung in Form des u_h^* vorliegt. Das Defektkorrektur-Prinzip kann iterativ angewandt werden, man erhält dann iterative Defektkorrektur-Verfahren (vgl. [5], [3]). Allerdings wird die Konstruktion der defektdefinierenden Operatoren $F_h^2, F_h^3, ...$ zunehmend aufwendiger.

Zur Absicherung der Existenz des v_h^* und der gewünschten Konvergenzeigenschaften sind folgende Banach-Räume $B^1, B^{0,1}$ glatter Funktionen einzuführen:

$$\begin{aligned} B^{0,1} &= C^1(\mathbb{T}^2, \mathbb{R}^q) \quad \text{mit Norm } \|u\|_1 \\ B^1 &= B \cap C^2(\mathbb{T}^2, \mathbb{R}^q) \end{aligned}$$

mit Norm

$$\|u\|_2 = max\{\|u\|_0, \|u_t\|_0, \|u_\theta\|_0, \|u_{tt}\|_0, \|u_{t\theta}\|_0, \|u_{\theta,\theta}\|_0\}.$$

Damit ist $B^1 \subset B$ und $B^{0,1} \subset B^0$. Entsprechende diskrete Analoga sind $B_h^1, B_h^{0,1}$ mit

$$\begin{aligned} B_h^{0,1} &= C_h^1(\mathbb{T}_h^2, \mathbb{R}^q) \quad \text{mit Norm } \|u_h\|_1 \\ B_h^1 &= B_h \cap C_h^2(\mathbb{T}_h^2, \mathbb{R}^q) \end{aligned}$$

mit der diskreten C^2-Norm $\|u_h\|_2$. Gemäß (10) werden darin die zweiten dividierten Differenzen $\partial_{tt}u_h, \partial_{t\theta}u_h, \partial_{\theta,\theta}u_h$ gebildet. Auch hierfür gelten die Inklusionen $B_h^1 \subset B_h,\ B_h^{0,1} \subset B_h^0$.

Sei $S^k(y,R) = \{z \in B^k | \|z-y\|_B^k \leq R\}$ eine Kugel in B^k, $k = 0, 1$, so findet man in [5], S. 62 ff, folgenden allgemeinen

Satz 5 *Für $k = 0, 1$ und konstantes $R > 0$ seien folgende Voraussetzungen erfüllt:*

V_1: *$Fu = 0$ ist isoliert lösbar in B^1 mit $u^* \in B^1$, eindeutig in $S^0(u^*, R)$.*

V_2: *F und F_h sind Fréchet-differenzierbar auf $S^0(u^*, R)$ bzw. $S^0(p_h u^*, R)$.*

V_3: *$F_h'(p_h u^*)$ besitzt einen beschränkten inversen Operator mit konstanten $S_k > 0$*

$$\|F_h'(p_h u^*)^{-1} v_h\|_{B_h^k} \leq S_k \|v_h\|_{B_h^{o,k}}, \quad k = 0, 1.$$

V_4: *F_h' ist gleichmäßig Lipschitz-stetig bzgl. h, d.h.*

$$\|F_h'(u_h) - F_h'(v_h)\|_{B_h^k, B_h^{o,k}} \leq L_k \|u_h - v_h\|_{B_h^k}, \quad k = 0, 1.$$

V_5: *F_h ist konsistent auf u^* mit Ordnung $p \in \mathbb{N}$, d.h.*

$$\|F_h p_h u^* - p_h F u^*\|_{B_h^{o,1}} \leq C \cdot h^p, \quad C > 0, \text{ konstant.}$$

V_6: *$F_h^1 : B_h^1 \to B_h^0$ ist Fréchet-differenzierbar auf $S^1(p_h u^*, R)$ mit*

$$\|F_h^{'}(u_h) - F_h^1(u_h)\| \leq C_{1,0} \cdot h^p \ , \ C_{1,0} > 0 \ , \ \textit{konstant.}$$

V_7: *F_h^1 ist konsistent mit Ordnung $2p$ auf u^* , d.h.*

$$\|F_h^1 p_h u^* - p_h F u^*\|_{B_h^0} \leq C_{2,0} \cdot h^{2p} \ , \ C_{2,0} > 0 \ , \ \textit{konstant.}$$

Dann existierten Konstanten $h^* > 0$ und $r_0 > 0, r_1 > 0$, so daß für alle Schrittweiten $0 < h \leq h^*$ gilt:

B_1: *Das allgemeine DC-Verfahren besitzt Lösungen, eindeutig in den Kugeln $u_h^* \in S^0(p_h u^*, r_0)$, $v_h^* \in S^1(p_h u^*, r_1)$.*

B_2: *u_h^* bzw. v_h^* konvergieren mit den Ordnungen p bzw. $2p$ gegen u^*, d. h.*

$$\begin{aligned} \|u_h^* - p_h u^*\|_{B_h^0} &\leq C_0 \cdot h^p, \\ \|v_h^* - p_h u^*\|_{B_h^0} &\leq C_1 \cdot h^{2p}. \qquad \square \end{aligned} \tag{40}$$

Wird das diskrete Problem (13) mit einem Newton-ähnlichen Verfahren gelöst, so wird man dies auch mit der korrigierten Gleichung (39) tun, womit sich die Folge $\{v_h^m\}$ mit

$$\begin{aligned} F_h^{'}(v_h^m)(v_h^{m+1} - v_h^m) &= -F_h v_h^m + l_h, \\ & \quad m = 0, 1, 2, ... \end{aligned} \tag{41}$$

ergibt. Sinnvoll ist der Startwert $v_h^0 = u_h^*$. In [5] wird ein konstruktives DC-Verfahren begründet, das mit einem einzigen Newtonschritt auskommt:

Algorithmus 3 (Konstruktive Defektkorrektur)

S_1: Lösung von $F_h u_h = 0$ mit Basislösung u_h^*.

S_2: Berechnung des Defektes $l_h = -F_h^1 u_h^*$.

S_3: Lösung der linearen Gleichung

$$F_h^{'}(u_h^*) e_h = l_h \ . \tag{42}$$

S_4: Bestimmung der Korrektur

$$v_h^* := u_h^* + l_h \ . \tag{43}$$

Es gilt folgender (vgl. [5]], S. 88, S. 95)

Satz 6 *Unter den Voraussetzungen V_1 bis V_7 des Satzes 5 gelten für u_h^* und das durch (43) ermittelte v_h^* die Behauptungen B_1 und B_2 des Satzes 5.*
Weiterhin gilt Behauptung

B_3: *Das durch (42) ermittelte e_h ist eine asymptotische Schätzung des globalen Diskretisierungsfehlers von u_h^*, d. h.*

$$e_h \quad = \quad u_h^* - p_h u^* + O(h^{2p}). \qquad \square \tag{44}$$

Bemerkungen:

1. Wird in (42) die Jacobimatrix $F_h^{'}(u_h^*)$ geeignet durch eine Matrix A mit

$$\|A - F_h^{'}(u_h^*)\|_{B_h, B_h^0} \quad \leq \quad D_1 \cdot h^p, \tag{45}$$

   $D_1 > 0$, konstant, approximiert, so gilt Satz 6 gleichfalls.

2. Benutzt man zur Lösung der korrigierten Gleichung (39) die Picardsche Fixpunktiteration (z. B. im Zusammenhang mit einem Schießverfahren zur Ermittlung der Anfangswerte $u(\theta, 0)$), so ist Satz 6 nicht anwendbar; die Lösung v_h^* von (39) muß im allgemeinen mit mehreren Iterationen approximiert werden.

3. In jedem Fall stellt

$$e_h \quad := \quad u_h^* - v_h^* \tag{46}$$

   eine asymptotische Schätzung des globalen Fehlers von u_h^* dar, denn

$$\begin{aligned} e_h \quad &= \quad (u_h^* - p_h u^*) - (v_h^* - p_h u^*) \\ &= \quad u_h^* - p_h u^* + O(h^{2p}) \end{aligned}$$

   gilt wegen der Behauptung B_2 obiger Sätze.

5 Defektoperator für 6-Punkt-Verfahren

Es wird angenommen, daß die Funktionen $\omega : \mathbb{T}^2 \times \mathbb{R}^q \to \mathbb{R}$ und $f : \mathbb{T}^2 \times \mathbb{R}^q \to \mathbb{R}^q$ hinreichend glatt sind und das Ausgangsproblem (6) eine isolierte (reguläre) Lösung $u = u^*(\theta, t)$ auf $\mathbb{T}^2$ besitzt. Für die 6-Punkt-Verfahren (12) gelten die Konsistenzbedingungen aus [1], wobei $\lambda := \tau/h = const.$ ist. Damit lassen sich die Voraussetzungen V_1 bis V_5 der obigen Sätze 5 und 6 verfizieren; die Konsistenzordnung ist $p = 1$.

Entscheidend für die Durchführbarkeit der DC-Verfahren 2 und 3 ist nun die Konstruktion des Defektoperators F_h^1, der die Voraussetzungen V_6 und V_7 des Satzes 5 erfüllt. Auf eine umfangreiche Herleitung des F_h^1 soll hier verzichtet werden. Wir definieren vielmehr F_h^1 und weisen die geforderten 2 Eigenschaften nach.

Satz 7 *Für $u_h = \{u_j^n\} \in B_h^1$ sei F_h^1 durch*

$$\begin{aligned}\{F_h^1 u_h\}_j^n \;\equiv\; & \frac{1}{\tau}[u_j^{n+1} - u_j^n] \\ & -\frac{1}{2}[f(\theta_j, t_n, u_j^n) - \omega(\theta_j, t_n, u_j^n)\frac{1}{2h}(u_{j+1}^n - u_{j-1}^n)] \\ & -\frac{1}{2}[f(\theta_j, t_{n+1}, u_j^{n+1}) - \omega(\theta_j, t_{n+1}, u_j^{n+1})\frac{1}{2h}(u_{j+1}^{n+1} - u_{j-1}^{n+1})], \\ & \textit{mit } j = 0(1)J-1, \quad n = 0(1)N-1\end{aligned} \tag{47}$$

definiert. Dann gelten die Behauptungen der Sätze 5 und 6 mit $v_h^* = \{v_h^*\}_j^n$, $u_h^* = \{u_h^*\}_j^n$, d.h.

$$v_j^n = u^*(\theta_j, t_n) + O(h^2), \tag{48}$$

und

$$e_j^n = u_j^n - v_j^n \tag{49}$$

ist eine asymptotische Schätzung des globalen Fehlers von u_h^*. □

Beweis: Der Nachweis der Voraussetzung V_7 ist leicht zu führen: Sei u^* die exakte Lösung, womit folgende Abkürzungen eingeführt werden:

$$\begin{aligned} u_j^n &:= u^*(\theta_j, t_n) \\ f_j^n &:= f(\theta_j, t_n, u_j^n) \\ \omega_j^n &:= \omega(\theta_j, t_n, u_j^n). \end{aligned} \tag{50}$$

Für die Differenz

$$d := F_n^1 p_n u^* - p_n F_n^*$$

erhält man punktweise mit (47)

$$\begin{aligned} d_j^n \;=\; & \frac{1}{\tau}[u_j^{u+1} - u_j^n] \\ & -\frac{1}{2}[f_j^n - \omega_j^n \cdot \frac{1}{2h}(u_{j+1}^n - u_{j-1}^n)] \\ & -\frac{1}{2}[f_j^{n+1} - \omega_j^{n+1} \cdot \frac{1}{2h}(u_{j+1}^{n+1} - u_{j-1}^{n+1})]. \end{aligned}$$

Da die zentralen Differenzen die Ableitung u_θ mit Ordnung $O(h^2)$ approximieren, erhält man

$$\begin{aligned} d_j^n \;=\; & \frac{1}{\tau}[u_j^{n+1} - u_j^u] \\ & -\frac{1}{2}\left[f_j^n - \omega_j^n \cdot \frac{\partial u}{\partial \theta}\Big|_j^n\right] - \frac{1}{2}\left[f_j^{n+1} - \omega_j^{n+1} \cdot \frac{\partial u}{\partial \theta}\Big|_j^{n+1}\right] + O(h^2), \end{aligned}$$

woraus sich mit DGl (6) an den Punkten (θ_j, t_n) und (θ_j, t_{n+1})

$$d_j^n = \frac{1}{\tau}[u_j^{n+1} - u_j^n] - \frac{1}{2}\left[\frac{\partial u}{\partial t}\Big|_j^n + \frac{\partial u}{\partial t}\Big|_j^{n+1}\right] + O(h^2)$$

ergibt. Taylorentwicklung liefert unmittelbar $d_j^n = O(\tau^2) + O(h^2)$, woraus mit $\lambda := \tau/h = const.$ die Voraussetzung V_7 folgt.
Der Nachweis der Approximationseigenschaft V_6 ist im allgemeinen Fall der Verfahrensklasse (12) ungleich aufwendiger: Seien $u_h,\ v_h \in B_h^1$ mit zugehörigen Normen $\|u_h\|_2,\ \|v_h\|_2$, die die Funktionen sowie die ersten und zweiten dividierten Differenzen enthalten. Da u_h und v_h Gitterfunktionen sind, sind keine Taylorentwicklungen möglich. Für die Operatorableitungen $F_h^{'}(u_h)v_h$ und $F_h^{1'}u_h v_h$ erhält man an der Stelle (θ_j, t_n)

$$\begin{aligned}
\{F_h^{'}(u_h)v_h\}_j^n = \frac{1}{\tau}\{ & \sum_{\mu=-1}^{1} \frac{\partial S_\mu^*}{\partial u}(\theta_j, t_n, u_j^n) v_j^n u_{j+\mu}^{n+1} \\
& + \sum_{\mu=-1}^{1} S_\mu^*(\theta_j, t_n, u_j^n) v_{j+\mu}^{n+1} \\
& - \sum_{\mu=-1}^{1} \frac{\partial S_\mu}{\partial u}(\theta_j, t_n, u_j^n) v_j^n u_{j+\mu}^{n} \\
& - \sum_{\mu=-1}^{1} S_\mu(\theta_j, t_n, u_j^n) v_{j+\mu}^{n}\} \\
& - \frac{\partial f}{\partial u}(\theta_j, t_n, u_j^n) v_j^n
\end{aligned} \tag{51}$$

sowie

$$\begin{aligned}
\{F^{1'}(u_h)v_h\}_j^n = & \frac{1}{\tau}[v_j^{n+1} - v_j^n] \\
& - \frac{1}{2}\{\frac{\partial f}{\partial u}|_j^n v_j^n - \omega_j^n \frac{1}{2h}[v_{j+1}^n - v_{j-1}^n] \\
& \quad - \frac{\partial \omega}{\partial u}|_j^n v_j^n \cdot \frac{1}{2h}[u_{j+1}^n - u_{j-1}^n]\} \\
& - \frac{1}{2}\{\frac{\partial f}{\partial u}|_j^{n+1} v_j^{n+1} - \omega_j^{n+1} \frac{1}{2h}[v_{j+1}^{n+1} - v_{j-1}^{n+1}] \\
& \quad - \frac{\partial \omega}{\partial u}|_j^{n+1} v_j^{n+1} \cdot \frac{1}{2h}[u_{j+1}^{n+1} - u_{j-1}^{n+1}]\} \,.
\end{aligned} \tag{52}$$

Durch Einfügen von Zwischentermen und Nutzung der Beziehungen zwischen den Matrizen S_μ^* und S_μ gelingt es, für die Differenz

$$\delta_j^n := \{F_h^{'}(u_h)v_h\}_j^n - \{F_h^{1'}(u_h)v_h\}_j^n$$

eine Abschätzung der Form

$$|\delta_j^n| \leq C_{1,0} \cdot h \cdot \|v_h\|_2 \qquad \forall_{j,n}$$

zu etablieren, deren rechte Seite den Faktor h (bzw. τ) enthält. Durch Maximumbildung erhält man hieraus

$$\|F_h^{'}(u_h)v_h - F_h^{1'}(u_h)v_h\|_{B_h^0} \leq C_{1,0} \cdot h \cdot \|v_h\|_{B_h^1}\ , \tag{53}$$

womit sich Voraussetzung V_6 ergibt. □

6 Algorithmische Umsetzung und Anwendungen

Aus Effizienzgründen benutzen wir das konstruktive DC-Verfahren entsprechend Algorithmus 3, das in einem ersten Schritt die Basislösung $u_h^* = \{u_h^*\}_j^n,\quad j = 0(1)J-1,\ \ u = 0(1)N-1$, mit einem Newton-ähnlichen Verfahren bestimmt. Das geschieht durch iterative Lösung von Anfangswertaufgaben mittels der Schießmethode. Im Anschluß sind nun 2 Vorgehensweisen möglich: Liegt die Lösung u_h^* gespeichert auf ganz $\mathbb{T}_h^2$ vor, so kann für jeden Gitterpunkt (θ_j, t_n) der Defektvektor $\{l_h\}_j^n$ mit

$$l_h = -F_h^1 u_h^*$$

gemäß (47) ermittelt werden und anschließend das modifizierte Problem

$$\begin{aligned} \frac{1}{\tau}\ \{\ & \sum_{\mu=-1}^{1} S_\mu^*(\theta_j, t_n, v_j^n) v_{j+\mu}^{n+1} - \\ & \sum_{\mu=-1}^{1} S_\mu(\theta_j, t_n, v_j^n) v_{j+\mu}^n\} - f(\theta_j, t_n, v_j^n)\ =\ l_j^n, \\ & v_j^N - v_j^0 = 0, \quad j = 0(1)J-1 \end{aligned} \tag{54}$$

in einem dritten Schritt gelöst werden. Wegen der Speicherung von u_h^* und l_h ist dieses serielle Vorgehen jedoch ineffizient. Günstiger erscheint ein zeitschrittweises Vorgehen, bei dem lediglich die Anfangswerte u_j^0 von u_h^*, $j = 0(1)J-1$, gespeichert werden. Definiert man in (54) die modifizierten rechten Seiten zu

$$\overline{f}_j^n := f(\theta_j, t_n, v_j^n) + l_j^n, \quad (\theta_j, t_n) \in \mathbb{T}_h^2, \tag{55}$$

so läßt sich (54) mit demselben Schießverfahren wie die Basisgleichungen lösen. Man bestimmt also im n-ten Zeitschritt durch nochmalige Integration der Basisgleichungen u_j^{n+1} und errechnet mit u_j^n und n_j^{n+1}, $j = 0(1)J-1$, den Defekt l_j^n und die modifizierten rechten Seiten $\overline{f}_j^n$. Damit kann der n-te Zeitschritt von (54) ausgeführt werden etc.

Nachfolgend soll der Effekt des konstruktiven DC-Verfahrens entsprechend Algorithmus 3 an zwei Beispielen demonstriert werden. Die globale Extrapolation dagegen erwies sich im allgemeinen als unterlegen; deshalb soll auf entsprechende Beispiele verzichtet werden.

Beispiel 1 Die skalare DGL auf dem 2-Torus

$$\frac{\partial u}{\partial t} + \omega \frac{\partial u}{\partial \theta} \;=\; u(s - u^2)\;, \quad (\theta, t) \in \mathbb{T}^2 \tag{56}$$

mit Konstanten $\omega = 0.57$ und $s = 0.16$ besitzt die Nullösung $u^*(\theta, t) = 0$ und die konstante Lösung $u^*(\theta, t) = \sqrt{s} = 0.4$, die damit zu Vergleichszwecken verfügbar ist. Die Rechnung wurde mit 3 *Basisverfahren* und der geforderten Genauigkeit TOL durchgeführt und lieferte folgende maximale absolute Fehler $\|u - u^*\|$ auf $\mathbb{T}_h^2$ (J bzw. N ist die Anzahl der Teilintervalle in $\theta-$Richtung bzw. in $t-$Richtung):

Tabelle 1: Maximaler absoluter Fehler bei Basis- und DC-Verfahren

Basisverfahren	TOL	J	N	$\|u - u^*\|$	$\|v - u^*\|$
Upwind, explizit	1E-4	20	80	5.280E-7	8.101E-9
CIR, explizit	1E-3	20	80	6.299E-5	5.883E-6
Upwind, explizit	1E-3	20	80	3.805E-5	2.945E-6
Upwind, implizit	1E-3	20	40	7.278E-5	7.179E-6

Anwendung des Algorithmus 3 ergab mit einem zusätzlichen Newtonschritt die maximalen Fehler $\|v - u^*\|$ der korrigierten Lösung v , die damit stets um eine Dezimalstelle genauer als die Basislösung war.

Beispiel 2 Nichtlineare parametrisch erregte elektrische Netzwerke lassen sich durch heteronome Differentialgleichungen (vgl. [4])

$$\ddot{x} + \alpha \dot{x}^3 - \beta \dot{x} + (1 + B \sin 2t)x \;=\; 0 \tag{57}$$

mit Parametern

$$\begin{aligned} B &= 0.1 \\ \alpha &= \varepsilon - B = \varepsilon - 0.1 \\ \beta &= \frac{\varepsilon}{2} - B = \frac{\varepsilon}{2} - 0.1 \end{aligned}$$

beschreiben. Nach Transformation in Polarkoordinaten erhält man das DGL-System in Radius-Winkel-Koordinaten

$$
\begin{aligned}
\frac{d\theta_1}{dt} &= \beta cs - sin^2\theta_1 - \alpha u^2 \sin^3\theta_1 \cos\theta_1 - \\
&\quad -(1 + B\sin 2\theta_2)\cos^2\theta_1 = \psi(\theta_1, \theta_2, u, \varepsilon) \\
\frac{d\theta_2}{dt} &= 1 \\
\frac{du}{dt} &= u(cs + \beta\sin^2\theta_1) - \alpha u^3 \sin^4\theta_1 - \\
&\quad -u(1 + B\sin 2\theta_2)cs = r(\theta_1, \theta_2, u, \varepsilon)
\end{aligned}
\tag{58}
$$

mit $cs = \cos\theta_1 \sin\theta_1$. Die Gleichung für den invarianten Torus wird folglich (vgl. [6]) durch die quasilineare DGL auf dem 2-Torus

$$
\frac{\partial u}{\partial \theta_2} + \psi(\theta_1, \theta_2, u, \varepsilon)\frac{\partial u}{\partial \theta_1} = r(\theta_1, \theta_2, u, \varepsilon)\,, \quad (\theta_1, \theta_2) \in \mathbb{T}^2 \tag{59}
$$

beschrieben. Für den Parameterwert $\varepsilon = 5.0$ wurde eine Referenzlösung mit dem expliziten Upwind-Verfahren und Picard-Iteration auf dem feinen $160 \times 320-$Punkt-Gitter ermittelt. Tabelle 2 gibt die ermittelten Näherungswerte $u(0,0)$ für die exakten Lösungen $u^*(0,0)$ an. Die mit den DC-Algorithmen 2 und 3 verbesserten Lösungen $v(0,0)$ sind in der letzten Spalte der Tabelle dargestellt.

Tabelle 2: Näherungslösungen bei Basis- und DC-Verfahren

Verfahren	J	N	$u(0,0)$	$v(0,0)$
Explizit Upwind + Picard	160	320	1.170 748	—
Implizit Upwind + Picard	160	320	1.168 294	—
Explizit Upwind	40	80	1.099 639	1.158 373
+ Newton-Verfahren	40	160	1.098 574	1.157 632
+ DC mit 3 Iterationen	50	100	1.113 998	1.171 668
Explizit Upwind + Newton	40	80	1.099 639	1.158 372
+ konstruktive DC	40	160	1.098 574	1.157 632
Implizit Upwind + Newton	40	80	1.095 686	1.154 464
+ konstruktive DC	50	100	1.110 258	1.168 147

Insbesondere das implizite Upwind-Verfahren mit konstruktiver Defektkorrektur gemäß Algorithmus 3 lieferte auf dem $50 \times 100-$Punkt-Gitter Lösungen mit derselben Genauigkeit - nur waren statt der 51200 Unbekannten nun lediglich 5000 Lösungswerte zu bestimmen!

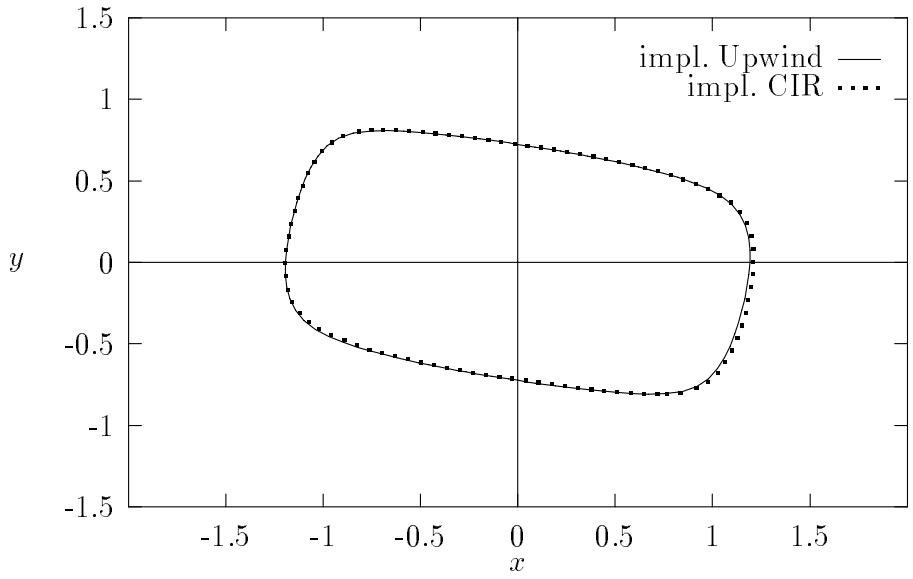

Abbildung 1: Torusquerschnitte für $\varepsilon = 5.0$

In Abbildung 1 sind die Torusquerschnitte für $\theta_2 = 0$ im cartesischen Koordinatensystem mit x und $y = \dot{x}$ dargestellt. Im Vergleich des (i) impliziten CIR-Schemas und anschließender Defektkorrektur mit dem (ii) impliziten glatten Upwind-Schema und anschließender Defektkorrektur wird eine sehr gute Übereinstimmung beider Kurven sichtbar. Der geschätzte asymptotische Fehler betrug beim CIR-Schema 0.0245 und beim impliziten glatten Upwind-Schema nur 0.007.

Literatur

[1] Bernet, K.; Vogt, W.: Anwendung finiter Differenzenverfahren zur direkten Bestimmung invarianter Tori. ZAMM 74 (1994), No. 6, T577-T579.

[2] Nowak, U.: Adaptive Linienmethoden für nichtlineare parabolische Systeme in einer Raumdimension. TR 93-14, Dez. 1993, ZIB.

[3] Böhmer, K.; Stetter, H.J. (Hrsg.): Defect Correction Methods - Theory and Applications. Springer-Verlag, Wien 1984.

[4] Philippow, E.S.; Büntig, W.G.: Analyse nichtlinearer dynamischer Systeme der Elektrotechnik. Carl Hanser Verlag, München-Wien 1992.

[5] Vogt, W.: Zur Theorie und Anwendung konstruktiver Defektkorrektur-Verfahren. Dissertation (B) an der TH Ilmenau, 1984, 144 S.

[6] Vogt, W.; Bernet, K.: A Shooting Method for Invariant Tori. Preprint No. M 3/95, Technical University of Ilmenau, Department of Mathematics.